BEBR
FACULTY WORKING PAPER NO. 89-1617

The Good, The Bad and The Ugly: Coalition Proof Equilibrium in Games with Infinite Strategy Spaces

Charles M. Kahn
Dilip Mookherjee

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois Urbana-Champaign

FACULTY WORKING PAPER NO. 89-1617

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

December 1989

The Good, The Bad and the Ugly: Coalition Proof Equilibrium in Games with Infinite Strategy Spaces

Charles M. Kahn
University of Illinois, Urbana-Champaign

Dilip Mookherjee
Indian Statistical Institute

We thank William Thomson and Carlos Arilis for useful comments, as well as participants at the summer game theory conferences at Ohio State. This work was funded by NSF grant SES-8821723.

Addresses: Kahn: Department of Economics, University of Illinois, 1206 S. Sixth Street, Champaign, IL 61820, USA. Mookherjee: Indian Statistical Institute 7 S.J.S. Sansanwal Marg, New Delhi, 110016, INDIA

Abstract:

This paper shows how to extend the definition of Coalition Proof Nash Equilibrium to games with infinite strategies. Our new definition reduces to the recursive definition of Bernheim Peleg and Whinston when there are a finite number of players and a finite strategy space. Unlike the old characterization, our new one maintains an equivalence between the recursive definition and a non-recursive characterization in terms of consistent families of equilibria; for the old definition the equivalence was only maintained in the case of finite players and finite strategy space. We give some examples of the employment of the new definition to show its advantages and its relatively simple characterization, even for games with infinite numbers of agents.

Mail correspondence to:

Charles M. Kahn
Department of Economics
University of Illinois
1206 S. Sixth St.
Champaign, IL 61820

THE GOOD, THE BAD AND THE UGLY:

COALITION PROOF EQUILIBRIUM IN GAMES WITH INFINITE STRATEGY SPACES

I. Introduction

In non-cooperative games with the pre-play communication, Bernheim-Peleg-Whinston (B-P-W) have argued that one should focus attention on a refinement of Nash equilibrium, called Coalition Proof Nash Equilibrium (CPNE). This concept appears a promising way to model the ability of coalitions of economic agents to co-ordinate their activities in an incentive compatible fashion. The B-P-W definition is complicated, requiring the use of a recursion on the number of players considered in the game. Moreover, we show below that the B-P-W definition appears to provide unreasonable solutions in games with infinite strategy spaces.[1] Since most economic applications involve infinite strategy spaces, a modification of the solution concept is called for.

There is an alternative characterization of Coalition Proof Equilibrium in terms of consistency of a set of equilibria for a family of games (see for example Greenberg [1986]). Surprisingly, however, the consistency characterization is <u>not</u> equivalent to the B-P-W definition in the case of games with infinite strategy spaces. We provide an example where the consistency characterization yields no equilibria but there are equilibria according to the B-P-W definition.

We show that situations in which the correspondence fails are situations in which the consistent set does not exist. Therefore we provide a weakened notion which we call semi-consistency.[2] We prove that families of semi-consistent solutions always exist, and that the family is unique in

finite player games. The semi-consistent characterization yields a refinement of the B-P-W definition; moreover, as we show in examples, cases where they diverge are cases where the B-P-W definition yields unreasonable outcomes.

Still it is valuable to restore the correspondence between the recursive and the consistency definitions. In order to do this we further modify the solution concept to incorporate a kind of "near-rational" behavior for coalitions. With this modification, it turns out that the difficulties noted above disappear: There always exists a consistent set, and the equilibria so characterized are precisely the CPNE of the recursive definition.

Our method for modeling near-rationality is of interest in its own right: Rather than following the standard approach of epsilon equilibrium, we proceed by modeling agents as choosing, not single actions, but convergent sequences of actions, with payoffs defined by the limit point of these action sequences. The interpretation is that they choose actions "sufficiently far along" these sequences, and thereby gain little by going to the limit.

Finally we consider the case of games with infinite numbers of players. Since the consistency characterization does not depend on finite recursions on the set of players, it is the suitable candidate for extension to the infinite player case. We demonstrate that a semi-consistent partition always exists, but need not be unique. However we show that the <u>minimal</u> semi-consistent partition can be easily characterized, and therefore propose as the natural extension a "strong" coalition proof equilibrium, characterized as belonging to <u>every</u> semi-consistent partition.

A natural economic application for a game with an infinite number of players is any situation with free entry. For the use of the solution concepts of this paper to investigate a problem in information economics, see Kahn-Mookherjee [1989].

II. Simple Agreements

Let N be the (countable) set of players, with $i \in N$ denoting a typical player. Let A_i, U_i respectively denote the strategy space and payoff functions of i and let $A \equiv \times_{i \in N} A_i$.

$$U_i : A \rightarrow \mathbb{R}.$$

Define an <u>agreement</u> to be a pair $(\underline{a}, S)$ where $\underline{a} \in A$ and $S \subseteq N$. Let $\mathbb{A}$ denote the set of possible agreements.

Note that an agreement specifies actions for all players, not just those in S. The reader may find it helpful to interpret an agreement as specifying the actions for the parties to the agreement, <u>given</u> the actions of other players.

Agreement $(\underline{b}, S)$ <u>trumps</u> agreement $(\underline{a}, T)$ (denoted $(\underline{b}, S) \gg (\underline{a}, T)$) if

(i) $S \subseteq T$

(ii) $a_j = b_j$ for all $j \in N \backslash S$

(iii) $U_i(\underline{b}) > U_i(\underline{a})$ for all $i \in S$.

In other words, a trumping agreement is one in which a subset of the original parties break away and find actions which are strictly better for all of the subset, the rest of the players leaving their actions unchanged.

The pair (G, B) is a <u>consistent partition</u> of the set of agreements $\mathbb{A}$ -- where G denotes the set of "good" agreements, and B denotes its complement, the set of "bad" agreements -- if

$(\underline{a},T) \in B \iff$ There exists $(\underline{b},S) \in G$ such that $(\underline{b},S) \gg (\underline{a},T)$

$(\underline{a},T) \in G \iff$ There does not exist $(\underline{b},S) \in G$ such that $(\underline{b},S) \gg (\underline{a},T)$

In other words, every agreement in G is trumped only by agreements in B, and every agreement in B is trumped by some agreement in G.[3]

It is not clear that a consistent partition always exists, or if it does, whether it is unique. However, if a consistent partition {G,B} does exist, it generates a solution concept in the following way: the set of solutions to the game is the set of strategy vectors a such that (a,N) is in the set G.

In the case of a finite number of players and strategies, it can be shown that there always exists a unique partition; this is established below. But in finite-player, infinite-strategy games, a consistent partition may not exist.

Example 1 Consider a one player game $N = \{i\}$, $A_i = (0,1)$ and $U_i(a_i) = a_i$. If G is non-empty, it must contain a single agreement -- otherwise one agreement in G will trump another. Let $G = \{(a_i^*, \{i\})\}$. Then $(a_i^*+\epsilon, \{i\}) \in B$, and therefore must be trumped by some agreement in G, which is a contradiction. G cannot be empty, because then every agreement would be in B, and by definition every agreement in B must be trumped by some agreement in G.■

Example 1 would no longer hold if the player had a compact strategy space. But with more players, this problem may arise even with compact strategy spaces and continuous utility functions.

Example 2 Suppose $N = \{1,2,3\}$, $A_i = [0,1]$ for all i and

$$U_1(a_1,a_2,a_3) = a_2 - |a_1-a_2|$$
$$U_2(a_1,a_2,a_3) = 2a_1 - a_3 - |a_2-a_3|$$
$$U_3(a_1,a_2,a_3) = 2a_1 - a_2 - |a_1-a_3|(1-a_1+a_2).$$

The best response correspondences are:

$$\tilde{a}_1(a_2,a_3) = a_2$$
$$\tilde{a}_2(a_1,a_3) = a_3 \quad \text{and}$$
$$\tilde{a}_3(a_1,a_2) = \begin{cases} [0,1] & \text{if } a_1 = 1 \text{ and } a_2 = 0 \\ a_1 & \text{otherwise.} \end{cases}$$

Suppose there exists a consistent partition $\{G,B\}$. First note that $((a_1,a_2,a_3),\{i\}) \in G$ if $a_i \in \tilde{a}_i(\underline{a}_{-i})$, where $\underline{a}_{-i}$ denotes the strategy-tuple of players other than *i* -- this is because such agreements cannot be trumped, as i is playing a best response. Next, note that $(\underline{a},N) \in G$ implies $\underline{a}$ must be a Nash equilibrium -- otherwise it can be trumped by a singleton-agreement involving a best-response strategy. Hence $(\underline{a},N) \in G$ implies $a_1 = a_2 = a_3$.

We claim that there is no $(\underline{a},N) \in G$. Suppose $((a,a,a),N) \in G$ and $((b,b,b),N) \in G$ with $b > a$; then the latter would trump the former, a contradiction. Suppose there is a unique $a \in [0,1]$ such that $((a,a,a),N) \in G$. If $a < 1$, then it follows that $((a+\epsilon,a+\epsilon, a+\epsilon),N) \in B$, which requires $((a+\epsilon,a+\epsilon,a+\epsilon),N)$ to be trumped by some agreement in G. Clearly it cannot be trumped by any singleton coalition agreement, nor by $((a,a,a),N)$. So $((a+\epsilon,a+\epsilon,a+\epsilon),N)$ must be trumped by some pair-coalition agreement. Suppose it is trumped by $((\alpha,\beta,a+\epsilon),\{1,2\}) \in G$. For this it is necessary that α and β both exceed $a+\epsilon$. But if $\beta > a+\epsilon$, $((\alpha,\beta,a+\epsilon),\{1,2\})$ is trumped by $((\alpha,a+\epsilon,a+\epsilon),\{2\}) \in G$, contradicting the hypothesis that $((\alpha,\beta,a+\epsilon),\{1,2\}) \in G$. Suppose, alternatively, that $((a+\epsilon,a+\epsilon,a+\epsilon),N)$ is trumped by $((a+\epsilon,\beta,\gamma),\{2,3\})$

$\in G$. This requires $\beta = \gamma < a+\epsilon$, but $a+\epsilon < 1$ implies that $((a+\epsilon,\beta,\gamma),\{2,3\})$ is trumped by $((a+\epsilon,\beta,a+\epsilon),\{3\}) \in G$, a contradiction. Finally, it is straightforward to show that $((a+\epsilon,a+\epsilon,a+\epsilon),N)$ cannot be trumped by any agreement involving coalition $\{1,3\}$.

It remains to consider the possibility that $((1,1,1),N)$ is the unique agreement for N that is in G. But $((1,1,1),N)$ is trumped by $((1,0,0),\{2,3\}) \in G$.

Hence there exists no $(\underline{a},N) \in G$. For any $a < 1$, then, $((a,a,a),N)$ must be in B, and therefore trumped by some agreement in G. Repeating the same argument as for $((a+\epsilon,a+\epsilon,a+\epsilon),N)$, however, this can be ruled out.■

We now propose the following modification of the notion of a consistent partition. A <u>semi-consistent</u> <u>partition</u> $\{G,U,B\}$ of the set of agreements $\mathbb{A}$ -- into a good set G, an ugly set U, and a bad set B -- is one where

$$(\underline{a},T) \in B \Leftrightarrow \text{There exists } (\underline{b},S) \in G \text{ such that } (\underline{b},S) \gg (\underline{a},T)$$

$$(\underline{a},T) \in G \Leftrightarrow (\underline{b},S) \gg (\underline{a},T) \text{ only if } (\underline{b},S) \in B$$

$$U = \mathbb{A} \setminus (G \cup B)$$

In other words B consists of all agreements trumped by agreements in G, and G consists of all agreements which are not trumped, except possibly by agreements in B.[4]

A semi-consistent partition is weaker than a consistent partition in that the good and the bad sets do not exhaust the set of all agreements: there may be agreements that are neither good, nor bad. The following lemma establishes some properties of this ugly set.

Lemma 1: (a) Any $(\underline{a},S)$ in U is trumped by some $(\underline{b},T)$ in U. Hence U is either empty or infinite.

(b) In finite player, finite-strategy games, U is empty. Hence a semi-consistent partition is a consistent partition in such games.

(c) $(\underline{a},S) \in U$ cannot be trumped by any $(\underline{b},T) \in G$. Hence with compact strategy sets and continuous utility functions, $(\underline{a},N) \in U$ implies $\underline{a}$ is a Nash equilibrium.

Proof (a) $(\underline{a},S) \in U$ must be trumped by some agreement $(\underline{b},T)$, otherwise it would be in G. If $(\underline{b},T) \in G$ then $(\underline{a},S)$ would be in B instead. If $(\underline{a},S)$ is not trumped by another agreement in U, then $(\underline{b},T) \in B$. But then $(\underline{a},S) \in G$, a contradiction.

(b) Follows from (a), since $\mathbb{A}$ is finite in finite games.

(c) is obvious. ■

Ugly sets are therefore "open" in the sense of containing infinite sequences of agreements trumping one another, but none of which are trumped by a good agreement. With compact strategy spaces and continuous utility functions, they involve members of the coalition playing best responses -- i.e., they are self-enforcing in the strict non-cooperative sense.

Theorem 1: A semi-consistent partition always exists.

Proof Define:

$$G_0 = \{(\underline{a},S)\in\mathbb{A} \mid \text{there is no } (\underline{b},T) \text{ in } \mathbb{A} \text{ trumping } (\underline{a},S)\}$$

$$B_0 = \{(\underline{c},V)\in\mathbb{A} \mid \text{there is some } (\underline{a},S)\in G_0 \text{ trumping } (\underline{c},V)\}.$$

Now define G_i, B_i inductively for $i = 1,2,3, \cdots$, by:

$$G_i = \{(\underline{a},S)\in\mathbb{A} \mid (\underline{b},T) \text{ trumps } (\underline{a},S) \Rightarrow (\underline{b},T) \in B_{i-1}\}$$

$$B_i = \{(\underline{c},V)\in\mathbb{A} \mid \text{there is } (\underline{a},S)\in G_i \text{ trumping } (\underline{c},V)\}.$$

It is obvious that the sets G_i and B_i form an increasing sequence.

Define $G = \bigcup_{i=0}^{\infty} G_i$ and $B = \bigcup_{i=0}^{\infty} B_i$.

We claim that $G\cap B = \emptyset$. Suppose otherwise that $(\underline{a},S) \in G\cap B$. Then for some i: $(\underline{a},S) \in G_\ell \cap B_\ell$ for all $\ell \geq i$. Now $(\underline{a},S)\in B_\ell$ implies that there exists $(\underline{b},T)\in G_\ell$ that trumps $(\underline{a},S)$. Since $(\underline{a},S)\in G_\ell$ as well, it follows that $(\underline{b},T)\in B_{\ell-1}$. So there exists $(\underline{c},V)\in G_{\ell-1}$ that trumps $(\underline{b},T)$. Since $(\underline{b},T)\in G_\ell$, it is trumped only by agreements in $B_{\ell-1}$; therefore $(\underline{c},V) \in B_{\ell-1} \cap G_{\ell-1}$.

Repeating this argument, there must exist $(\underline{d},W) \in B_0 \cap G_0$, a contradiction.

If $U = \mathbb{A} - (G\cup B)$, we establish that $\{G,U,B\}$ is a semi-consistent partition. Suppose $(\underline{a},S)\in G$, and it is trumped by $(\underline{b},T)$. Suppose $(\underline{a},S)\in G_\ell$. Then $(\underline{b},T)\in B_{\ell-1}$, implying $(\underline{b},T)\in B$. On the other hand, if $(\underline{b},T)\in B$, let $(\underline{b},T)\in B_\ell$. Then it is trumped by $(\underline{a},S)\in G_\ell$, which must be in G. ■

The previous theorem gives a procedure for generating a semi-consistent partition. The next theorem establishes that in a game with a finite number of players there is no other semi-consistent partition.

<u>Theorem</u> <u>2:</u> For finite player games, the semi-consistent partition is unique.

<u>Proof:</u> See appendix.

Theorems 1 and 2 give rise to the following definition:

A strategy vector $\underline{a}$ is a Semi-Consistent Coalition Proof Equilibrium (S-CPNE) if $(\underline{a},N) \in G$ in the semi-consistent partition of agreements.

Corollary 1: Every strong equilibrium is an S-CPNE.

Proof: In the proof of theorem 1, the strategy vectors $\underline{a}$ such that $(\underline{a},N)$ is in G_0 are precisely the strong equilibria.

It should be kept in mind that the existence of a CPNE is a separate issue from the issue of existence of a consistent or semi-consistent partition. A semi-consistent partition always exists, but no CPNE -- of either variety -- may exist. In example 2, this is precisely the case: there is no $(\underline{a},N)$ in the good set of the semi-consistent partition. All Nash equilibria of the form $((a,a,a), N)$ with $a < 1$ are in the ugly set, while $((1,1,1), N)$ is in the set B.

Obviously, when a consistent partition exists, the equilibria it generates are identical to the S-CPNE. We now explore the connection of S-CPNE with the recursive definition of a CPNE due to B-P-W in the case of finite player games.

Recursive Definition of CPNE: For any singleton coalition $\{i\}$, define $\underline{a}$ to be optimal for $\{i\}$ if i is playing a best response in $\underline{a}$.

Having defined optimality for all coalitions of size (k-1) or less, define optimality for a coalition S of size k (≥ 2), as follows.

Say that $\underline{a}$ is self-enforcing for $\underline{S}$ if it is optimal for every $T \subset S$.

Say that $\underline{a}$ is <u>optimal for S</u> if it is self-enforcing for S, and there does not exist any $\underline{b}$ self-enforcing for S such that ($\underline{b}$,S) trumps ($\underline{a}$,S).

Finally, if N is finite, say that $\underline{a}$ is an <u>R (recursive)-CPNE</u> if $\underline{a}$ is optimal for N.

The following example illustrates that the semi-consistent partition is an improvement on the consistent partition. It shows that there are cases in which the S-CPNE coincides with the R-CPNE, but there is no consistent partition. Since the S-CPNE is a strong equilibrium in this example, it also shows that corollary 1 would fail to hold if we used consistency rather than semi-consistency as the basis of the definition.

<u>Example 3:</u> $N = \{1,2\}$, $A_1 = [-2,1]$, $A_2 = [-1,1)$,

$$U_1(a_1,a_2) = U_2(a_1,a_2) = a_1 a_2.$$

In this example, strategy vector (-2,-1) strictly Pareto dominates all other strategy vectors. Since it is self enforcing, it is the unique R-CPNE. It is also the unique S-CPNE, since ((2,1),N) belongs to G_0. However there is no consistent partition; for example, agreements of the form $((1,\ a_2),\{2\})$ are neither good nor bad, by an argument identical to that of example 1.

In this example, the equilibrium is intuitively plausible, and the two definitions coincide.[5] In general the S-CPNE are a subset of the R-CPNE, as the following theorem demonstrates:

<u>Theorem 3:</u> In a finite player game, if $\underline{a}$ is a S-CPNE, it is a R-CPNE, but the converse is not true.

Proof Let $(\underline{a},S) \in G$ in a semi-consistent partition. We use an inductive method to establish that $\underline{a}$ is optimal for S. This is obviously true for any singleton S. So suppose it is true for all coalitions of size not exceeding k-1, and let #S = k.

First, we establish that $\underline{a}$ is self-enforcing for S. If not, there exists $T \subset S$ for which $\underline{a}$ is not optimal. By the inductive hypothesis, $(\underline{a},T) \notin G$. Therefore it is trumped by $(\underline{b},V) \notin B$. But then $(\underline{b},V)$ trumps $(\underline{a},S)$. This contradicts $(\underline{a},S) \in G$.

Next, we show there cannot be any $\underline{d}$ which is self-enforcing for S such that $(\underline{d},S)$ trumps $(\underline{a},S)$. Suppose there is. Since $(\underline{a},S) \in G$, $(\underline{d},S)$ must be in B.

We claim that there exists $(\underline{e},S) \in G$ which trumps $(\underline{a},S)$, which would lead to a contradiction. Since $(\underline{d},S) \in B$, there exists $(\underline{f},T) \in G$ which trumps $(\underline{d},S)$. If $T \subset S$, the induction hypothesis implies $\underline{f}$ is optimal for T. Then $\underline{d}$ cannot be optimal for T, contradicting the hypothesis that $\underline{d}$ is self-enforcing for S. So we must have $(\underline{f},S) \in G$ which trumps $(\underline{d},S)$ and therefore also $(\underline{a},S)$. Putting $\underline{e} = \underline{f}$, we are done. That the converse is not true is established by the following example. ■

Example 4: $N = \{1,2\}$, $A_1 = \{0,1\}$, $A_2 = [0,1)$,

$$U_1(a_1,a_2) = U_2(a_1,a_2) = a_1a_2.$$

In this game, there is a unique Nash equilibrium (0,0). Since this is the only self-enforcing strategy vector for {1,2}, it is optimal for {1,2}, and therefore a R-CPNE. We claim that $(\underline{0},N) \notin G$. Suppose otherwise. Now $(\underline{0},N)$ is trumped by $(\underline{a},N)$ where $a_1a_2 > 0$. So it must be that $(\underline{a},N) \in B$, and there exists $(\underline{c},S) \in G$ trumping $(\underline{a},N)$. If S = N, then $(\underline{c},N)$ would also trump $(\underline{0},N)$, a contradiction. Clearly S must be {2} and $c_1 = 1$, $c_2 > a_2$. So $((1,c_2),\{2\}) \in G$.

This is trumped by $((1,c_2+\epsilon),\{2\})$, which must therefore be in B. Any agreement that trumps this must be $((1,a_2^*),\{2\})$ with $a_2^* > c_2+\epsilon$. Therefore, there is $((1,a_2^*),\{2\}) \in G$, with $a_2^* > c_2+\epsilon > c_2$ -- so it trumps $((1,c_2),\{2\}) \in G$, a contradiction. ■

The R-CPNE in this example appears unreasonable, since it has player 1 choosing a dominated strategy. If 2 conjectures that 1 will not play a dominated strategy, 2 should play some a_2 close to 1. The importance of self enforcing agreements is that they protect a player from being "double-crossed" but in this example, they prevent players from enjoying mutually beneficial actions, solely because there is no "best" such action. Were we to allow in some fashion "near-rational" behavior for the coalition $\{1,2\}$, we would be able to describe actions $\underline{a}$ with a_1 and a_2 both close to 1 as "almost" self-enforcing, and therefore superior to $\underline{0}$. In the following section, we extend our approach to allow for such forms of "almost" self-enforcing agreements.

With a slightly more complicated example, we can show that R-CPNE need not be S-CPNE, even in a game with compact strategies and continuous utility functions:

<u>Example 5</u>:

$$N = \{1,2,3,4\},\ A_i = [0,1]\times\{0,1\}$$

$$U_1(\underline{a}) = p_1p_2p_3p_4[x_2 - |x_1-x_2|]$$

$$U_2(\underline{a}) = p_1p_2p_3p_4[2x_1 - x_3 - |x_2-x_3|]$$

$$U_3(\underline{a}) = p_1p_2p_3p_4[2x_1 - x_2 - |x_1-x_3|(1-x_1+x_2)]$$

$$U_4(\underline{a}) = p_1p_2p_3p_4[x_4]$$

where x_i denotes the real number in $[0,1]$ chosen by i, and $p_i \in \{0,1\}$ the second component of i's decision. (Think of this game as an extension of the game in example 2: Call any agreement with $p_1^k = p_2^k = p_3^k = p_4^k = 1$ a "participative" agreement; if all four players agree, players 1-3 play the game in example 2; otherwise, all players receive 0.)

It is obvious that any strategy with $p_1^k = p_2^k = p_3^k = p_4^k = 0$ is self-enforcing for N. We shall establish that any such agreement is an R-CPNE but not a S-CPNE.

To show that any such strategy $\underline{a}$ is an R-CPNE, note that any trumping agreement must be participative. However, no participative $\underline{b}$ is self-enforcing for N. Suppose otherwise; then it must constitute a Nash equilibrium, and $x_1 = x_2 = x_3 = x$, say, while $x_4 = 1$. If $x < 1$, then any participative strategy $\underline{c}$ with real-valued components $(x+\epsilon, x+\epsilon, x+\epsilon, 1)$ is self-enforcing for $\{1,2,3\}$. This is established by arguments analogous to those in example 2: neither of the first three players can unilaterally deviate profitably, and neither can any pair-subcoalition of $\{1,2,3\}$ engineer a coordinated deviation that is self-enforcing. So $\underline{b}$ cannot be self-enforcing for N, as $(\underline{b},N)$ is trumped by $(\underline{c},\{1,2,3\})$, implying that $\underline{b}$ cannot be optimal for $\{1,2,3\}$. If $x = 1$, then $(\underline{b},N)$ is trumped by $(\underline{d},\{2,3\})$, where $\underline{d}$ is a participative strategy with $x_2^k = x_3^k = 0$, which is optimal for $\{2,3\}$.

We next show that $(\underline{a},N) \notin G$ in any semi-consistent partition. Now $(\underline{a},N)$ is trumped by $(\underline{e},N)$ where $\underline{e}$ is a participative strategy vector and $x_1^k = x_2^k = x_3^k = x < 1$ and $x_4^k = 1$. If $(\underline{e}, N) \notin B$ then $(\underline{a},N) \notin G$, and we are done.

Therefore, suppose that $(\underline{e},N) \in B$. If so, there is something in G which trumps it. It cannot involve singletons, or paired coalitions, neither can it include player 4. So it must involve the coalition $\{1,2,3\}$. Further, to be in G, it must involve $x_1^k = x_2^k = x_3^k = x^*$ where $x < x^* < 1$. But any such agreement

must be in U, by reasoning which is identical to that of example 2. Therefore $(\underline{e},N) \in U$. Contradiction. ■

The two examples above establish that the recursive and non-recursive formulations are not identical in all cases. Since the equilibria generated by a consistent partition are S-CPNE, example 5 demonstrates a fortiori that the characterization in terms of consistent partitions is not equivalent to the recursive formulation. The following theorems establish circumstances in which the S-CPNE and the R-CPNE coincide:

<u>Theorem</u> <u>4</u> For finite player games, if a consistent partition exists the S-CPNE and R-CPNE coincide.

<u>Proof</u> Since a consistent partition is a semi-consistent partition with $U = \emptyset$, the uniqueness of the consistent partition follows from Theorem 2. Any agreement in B is trumped by an agreement in G and by Theorem 3 all agreements in G are optimal, so all agreements in B are non optimal. Since these two sets exhaust the set of agreements, the set G equals the set of optimal agreements. ■

In order to demonstrate the equivalence of the various versions of CPNE we require that the consistent partition exists. In one important case existence is easily established:

<u>Theorem</u> <u>5</u>: In any finite player, finite strategy game, a consistent partition exists.

<u>Proof</u>: By Theorem 1, a semi-consistent partition exists, and by Lemma 1

(a) the set U is empty. ■

Since our results seem to be different from those noted in Greenberg [1986] it is important to point the source of the discrepancy. Greenberg finds that R-CPNE can be characterized by using consistent partitions. His approach however <u>assumes</u> that a consistent partition exists. This cannot be demonstrated in general, except for finite-player, finite-action games.

III. Extended Agreements

The recursive and non-recursive definitions are equivalent for games with finite strategy spaces and finite numbers of players. The equivalence breaks down in other cases. In example 4 equivalence would be restored if we modified the game by adding strategies which corresponded to the limit points of the strategy space. But example 5 demonstrates that this proposed resolution will not work in general. In this section we will examine a more satisfactory resolution for the case of infinite strategy spaces.

For this section we will assume that the set of players N is finite, and that each player's utility function U_i is continuous on the compact strategy space $\times A_i$.

An <u>extended</u> <u>agreement</u> $(\underline{a}, S)$ consists of a coalition $S \subseteq N$ and a sequence of strategies $\underline{a} = (\underline{a}^1, \underline{a}^2, \ldots)$ which satisfies the following conditions:

1: The sequence converges.

2: For all $t \notin S$, $a_t^k = a_t^{k'}$ for all k, k'.

In other words, an extended agreement is a sequence of coordinated

actions by the members of coalition S, holding the actions of non-members fixed. Extended agreements include simple agreements as a special case: identify a simple agreement with an extended agreement in which the sequence of strategy vectors is constant. Let $\mathbb{A}^*$ denote the set of extended agreements.

An extended agreement ($\underline{b}$,T) <u>trumps</u> ($\underline{a}$,S) if

(a) $T \subseteq S$

(b) There exists k such that $b_j^k = a_j^k$ for all $j \in N\backslash T$

(c) $\lim U_i(\underline{b}^k) > \lim U_i(\underline{a}^k)$ for all $i \in T$.

It is useful to note that ($\underline{b}$,T) trumps ($\underline{a}$,S) implies that it also trumps ($\underline{a}$,V) for any $V \supseteq T$ if ($\underline{a}$,V) is in $\mathbb{A}^*$. Trumping agreements are coordinated deviations from initial agreements: At any step in the process that forms the initial extended agreement, say k, a subcoalition can agree to break away and follow their own coordinated deviation.[6]

The definition of a semi-consistent partition carries over without modification, as do the proofs of theorems 1 and 2 establishing the existence and uniqueness of the semi-consistent partition.[7]

A strategy vector $\underline{a}$* is an <u>Extended Semi-Consistent Coalition Proof Nash Equilibrium</u> (ES-CPNE) if it is the limit of a sequence of strategy vectors $\underline{a}$, where ($\underline{a}$,N) is in G for the semi-consistent partition of extended agreements.

For finite-player games, there is a recursive definition in extended agreements analogous to the recursive definition of the previous section:

<u>Recursive Definition of Extended CPNE:</u> For any singleton coalition {*i*}, say that $(\underline{a},\{i\}) \in \mathbb{A}^*$ is <u>optimal</u> if there does not exist any ($\underline{b}$, {i}) in $\mathbb{A}^*$ which trumps ($\underline{a}$, {i}).

Having defined optimality for all coalitions of size (k-1) or less, define optimality for a coalition S of size k ($\geq$2), as follows.

Say that ($\underline{a}$,S) $\in A^*$ is <u>self-enforcing</u> if there does not exist an optimal ($\underline{b}$,T) that trumps ($\underline{a}$,S), with T $\subset$ S.

Say that ($\underline{a}$,S) is <u>optimal</u> if it is self-enforcing and there does not exist any self-enforcing ($\underline{b}$,S) that trumps ($\underline{a}$,S).

Finally, if N is finite, say that the strategy vector $\underline{a}$* is an <u>Extended Recursive-CPNE</u> (ER-CPNE) if it is the limit of a sequence of strategy vectors $\underline{a}$ such that extended agreement ($\underline{a}$,N) is optimal.

The next result is the major result of the paper. It says that if we define coalition proof equilibria in terms of extended agreements, then for all finite player games -- those with infinite strategy spaces as well as finite strategy spaces -- the recursive and the non-recursive definitions are equivalent.

<u>Theorem</u> <u>6</u>: Using extended agreements, there is a unique semi-consistent partition, which is also a consistent partition. In this consistent partition, ($\underline{a}$,S) $\in$ G if and only if ($\underline{a}$,S) is optimal; hence the limit of $\underline{a}$ is an ER-CPNE if and only if it is an ES-CPNE.

<u>Proof:</u> See appendix.

A consequence of the proof of this theorem is that there is an equivalent, somewhat simpler recursive definition.

Corollary: The following definition is equivalent to the other definitions of Extended CPNE.

Alternative Recursive Definition of Extended CPNE: For any singleton coalition $\{i\}$, say that all $(\underline{a},\{i\}) \in \mathbb{A}^*$ are self-enforcing.

Having defined self-enforcing for all coalitions of size (k-1) or less, define it for a coalition S of size k ($k \geq 2$) as follows.

Say that $(\underline{a},S) \in \mathbb{A}^*$ is self-enforcing if there does not exist a self-enforcing $(\underline{b},T)$ that trumps $(\underline{a},S)$, with $T \subset S$.

Finally, if N is finite, say that the strategy vector $\underline{a}^*$ is an Extended-CPNE if it is the limit of a sequence of strategy vectors $\underline{a}$ such that the extended agreement $(\underline{a},N)$ is self-enforcing and not trumped by any self enforcing agreement.

Proof: See appendix.

It is instructive to use examples 2 and 5 to contrast the extended equilibrium definition with the initial definitions. In example 2, the unique extended equilibrium is the point (1,1,1). Although it is not self enforcing, it is the limit of a sequence of self enforcing agreements (x,x,x) and is the optimal member of this class. There was no R-CPNE and no S-CPNE.

Example 5 shows the power of our new definition. The fact that the set of optimal three player agreements is empty allows unreasonable R-CPNE in the four player game. Our definition of an S-CPNE eliminates these equilibria, but does not suggest an alternative. The extended equilibrium for this example is the point (1,1,1,1): it is "almost" self enforcing, and is maximal among such points.

IV. Games with a Countably Infinite Number of Players

Theorem 1, demonstrating the existence of a semi-consistent partition, does not depend on the number of players in the game being finite. But if there is a countable infinity of players, the semi-consistent partition need not be unique. This fact gives rise to the following definitions:

The strategy vector $\underline{a} \in A$ is a Weakly Consistent CPNE if $(\underline{a};N) \in G$ for some semi-consistent partition of $\mathbb{A}$. It is a Strongly Consistent CPNE if $(\underline{a},N) \in G$ for every semi-consistent partition.

The following example illustrates the distinction:

Example 6: $N = \{1,2,3\ldots\}$; $A_i = \{0,1\}$ for $i \in N$.

Thus a strategy vector is an infinite string of zeros and ones. In this game any strategy vector gives a payoff of zero to all players, *except* for strategy vectors listed in the table below. For a strategy vector $x^i = (x^i_1, x^i_2, x^i_3, \ldots)$ in the table, the corresponding payoff vector $u^i = (u^i_1, u^i_2, u^i_3, \ldots)$ is indicated to the right:

$x^1 = (1\ 1\ 1\ 1\ 1\ 1\ \ldots)$ $\qquad u^1 = (2,2,2,2,2,2,\ldots$

$x^2 = (1\ 0\ 0\ 0\ 0\ 0\ \ldots)$ $\qquad u^2 = (1,3,3,3,3,3,\ldots$

$x^3 = (1\ 0\ 1\ 1\ 1\ 1\ \ldots)$ $\qquad u^3 = (1,1,4,4,4,4,\ldots$

$x^4 = (1\ 0\ 1\ 0\ 0\ 0\ \ldots)$ $\qquad u^4 = (1,1,1,5,5,5,\ldots$

...

$x^\infty = (1\ 0\ 1\ 0\ 1\ 0\ \ldots)$ $\qquad u^\infty = (1,1,1,1,1,1,\ldots$

In this game there are two complete consistent partitions; in one all the even numbered strings are Weakly Consistent CPNE, in the other all the odd numbered strings are Weakly Consistent CPNE. There is no Strongly Consistent CPNE. Note that x^{∞} is not an equilibrium in either partition.■

It might seem a difficult matter to check whether an agreement is in every semi-consistent partition; in fact the procedure for doing so is relatively simple. This is due to the following theorem:

Theorem 7: There exists a minimal semi-consistent partition $\{G^*, U^*, B^*\}$ -- that is $\{G^*, U^*, B^*\}$ is a semi-consistent partition such that for every othersemi-consistent partition $\{G, U, B\}$ $G^* \subset G$, and $B^* \subset B$.

Proof: In the derivation in theorem 1, note that the set G_0 will belong to the G set in any semi-consistent partition, B_0 will belong to the B set in any semi-consistent partition, and if G_{i-1} belongs to any semi-consistent partition, then B_{i-1} and G_i must as well.■

Thus the minimal semi-consistent partition is precisely the semi-consistent partition generated in theorem 1. As an immediate consequence, we have the following characterization:

Corollary: $\underline{a}$ is a strongly consistent CPNE if and only if agreement $(\underline{a},N)$ belongs to some set G_i for i = 0,1,2.. where the sets G_i are defined as follows:

G_0 is the set of agreements which are not trumped by any agreement in $\mathbb{A}$.

For i = 1,2,.., G_i is the set of agreements which are trumped by no

agreement in $\mathbb{A}$, except for agreements which are in turn trumped by some agreement in G_{i-1}.

Note that this is a recursive characterization as well, but unlike the B-P-W definition, it does not require recursion on the number of members of the agreement. Thus this new characterization, unlike the B-P-W definition, is applicable to games with infinite numbers of players.

IV: Summary and Comments

In order to propose yet another modification of yet another solution concept, the proposer ought to give three sorts of justifications. First, to demonstrate that it is indeed a modification, the new definition and the original ought to be posed in such a form that the similarities are apparent. This we have done by comparing semi-consistent partitions with consistent partitions, and extended agreements with agreements.

Second, the proposer ought to show in examples that the new solution concept yields more "intuitive" results than does the original. We have done this as well: Compare the extended agreement with any of the initial agreements for example 5.

Third, and in our view most important, the proposer ought to demonstrate that theorems that hold for the original definition hold with greater generality or more regularity with the new definition. The result we have examined is the fact that Coalition Proof Equilibrium as recursively defined can be characterized by a consistency criterion. For the original definition this is true when action sets are finite, but not when they are infinite. For our definition with extended agreements the characterization holds generally.

When we extend the solution in extended agreements to situations with countably infinite numbers of players, semi-consistent partitions will no longer be unique, leading to weak and strong versions of the definition in terms of semi-consistency. In this case, we would argue that the strong version has the greater claim to our attention, if for no other reason than that theorem 1 gives us an explicit method for finding the solutions. This method can be given the following interpretation: An agreement is a good one if 1) <u>no</u> other agreement blocks it (strong equilibria would fall in this category) or if 2) no agreement blocks it except for agreements that are blocked by agreements of the first kind, or if 3) no agreement blocks it except agreements that are blocked by agreements of the first or second kind, and so forth.

Future papers will apply these two modifications to other solution concepts.

Appendix

In general trumping is not a transitive relation. Nonetheless, for both agreements and extended agreements the following is valid:

<u>Lemma</u> A.1: If $(\underline{c},S)$ trumps $(\underline{b},S)$ and $(\underline{b},S)$ trumps $(\underline{a},T)$ then $(\underline{c},S)$ trumps $(\underline{a},T)$.

<u>Proof:</u> Parts (i) and (iii) of the definition of trumping are immediately verified. Part (ii) obvious in the case of simple agreements; for extended agreements it is nearly so: if $\underline{b}_i^k = \underline{a}_i^k$ for $i \notin S$, then $\underline{c}_i^k = \underline{a}_i^k$ for $i \notin S$.■

<u>Note</u>: In particular, by setting $S = T$ we have transitivity on any set of agreements for a single coalition.

<u>Proof of Theorem 2:</u>

Consider any two semi-consistent partitions $\{G_1,U_1,B_1\}$ and $\{G_2,U_2,B_2\}$.

<u>Claim 1:</u> For any $i \in N$,

$(\underline{a}, \{i\}) \in G_1 \Leftrightarrow (\underline{a}, \{i\}) \in G_2$

and $(\underline{a}, \{i\}) \in B_1 \Leftrightarrow (\underline{a}, \{i\}) B_2$

<u>Proof:</u> If $(\underline{a},\{i\}) \in G_1$ then it cannot be trumped by any agreement. For any $(\underline{b},\{i\})$ which trumps it must be in B_1. But then $(\underline{b},\{i\})$ must be trumped by $(\underline{c},\{i\})$ in G_1 which then also trumps $(\underline{a},\{i\})$, a contradiction. Since $(\underline{a},\{i\})$ is not trumped by any agreement it must also be in G_2. Reversing the argument, $(\underline{a}, \{i\}) \in G_1 \Leftrightarrow (\underline{a}, \{i\}) \in G_2$.

Suppose $(\underline{a}, \{i\}) \in B_1$. Then there exists $(\underline{b}, \{i\})$ in G_1 which trumps it. Since $(\underline{b}, \{i\})$ must also be in G_2, it follows that $(\underline{a}, \{i\})$ is in B_2. Reversing the argument, the claim is established.

<u>Claim 2</u>: Suppose for all coalitions T with $\#T \leq k-1$, it is true that

$$(\underline{a}, T) \in G_1 \Leftrightarrow (\underline{a}, T) \in G_2$$

$$\text{and } (\underline{a}, T) \in B_1 \Leftrightarrow (\underline{a}, T)\ B_2$$

then the same is true for coalitions of size k.

<u>Proof</u>: Suppose for T of size k, $(\underline{a},T) \in G_1$, $(\underline{a},T) \notin G_2$. Then $(\underline{a},T)$ is trumped by $(\underline{b},V)$ which is not in B_2. If V is a proper subset of T, then the induction hypothesis implies that $(\underline{b},V)$ is not in B_1; contradicting $(\underline{a},T) \in G_1$.

The other possibility is that $V = T$, in which case, $(\underline{b},T)$ not in B_2 trumps $(\underline{a},T)$. Now $(\underline{a},T) \in G_1$ implies $(\underline{b},T) \in B_1$. So $(\underline{b},T)$ must be trumped by some $(\underline{c},W) \in G_1$. W cannot equal T; otherwise $(\underline{c},W)$ would also trump $(\underline{a},T)$. $W \subset T$ implies $(\underline{c},W) \in G_2$ implying $(\underline{b},T) \in B_2$, a contradiction. We conclude that $(\underline{a},T) \in G_1$ implies $(\underline{a},T) \in G_2$.

Suppose $(\underline{a},T) \in B_1$. Then there exists $(\underline{b},W) \in G_1$ which trumps it. By the above reasoning, $(\underline{b},W) \in G_2$, so $(\underline{a},T) \in B_2$.

Reversing the argument completes the proof. ■

Proof of Theorem 6:

The proof proceeds through a series of lemmas. We begin with a notational convention, some definitions, and a preliminary observation.

Notational Convention: If $\underline{d}$ is a convergent sequence of strategy vectors, we will use the notation $\underline{d}^*$ to represent the limit strategy vector.

For x, y in $\mathbb{R}^n$, let $x > y$ mean $x_i > y_i$ for $i = 1, .. n$. We will say set A dominates x if $a > x$ for some a in A. x is an upper bound for A if A does not dominate x.

Observation: If A is bounded and A dominates x, then there is a point z in the closure of A such that $z > x$ and z is an upper bound for A.

Proof of Observation: There exists a in A such that $a > x$. Define the set B to be points in the closure of A such that $m_i \geq a_i$ for $i = 1,..n$. B is a non-empty, compact set. Pick z in B to maximize $\Sigma_i z_i$. ■

Lemma A.2: For any non-empty, bounded set A in $\mathbb{R}^n$, either there is a point x in A which is an upper bound for A or there is a sequence of strictly increasing points in A whose limit is an upper bound for A.

Proof: Define a sequence recursively as follows:

Given a_{i-1}, if it is an upper bound for A, stop. Otherwise, using the above observation, pick $z_i > a_{i-1}$ which is an upper bound for A and which is the limit of a sequence of points in A. From that sequence, we can pick a_i such that a_i is strictly greater than a_{i-1} and the distance between a_i and z_i is less than 2^{-i}.

If the process terminates the theorem is proved. If the process does not

terminate, take a convergent subsequence of z_i's. The corresponding a_i's form the desired sequence. ■

The key result is the following lemma: it rules out "openness" of the set of self-enforcing agreements.

<u>Lemma A.3:</u> Suppose ($\underline{a}$, T) is not optimal, but is self enforcing (if #T $\geq$ 2) then there exists ($\underline{c}$, T) which is optimal and trumps ($\underline{a}$, T).

<u>Proof</u>: The result is obvious for a singleton coalition. For #T $\geq$ 2, let

$$\mathcal{C} = \{(\underline{d},T) \text{ self enforcing } \mid \underline{d}_i^k = \underline{a}_i^k \text{ for } i \notin T\}$$

In other words any self-enforcing agreement which trumps ($\underline{a}$,T) must be in $\mathcal{C}$. Consider the set Im $\mathcal{C}$, the image in $\mathbb{R}^{=T}$, the space of payoffs for members of T, from the limit strategy vectors for all extended agreements in $\mathcal{C}$. Note that an extended agreement in $\mathcal{C}$ is optimal if and only if its image is an upper bound in Im $\mathcal{C}$. Applying lemma A.2 to Im $\mathcal{C}$, we can conclude that either there is an optimal agreement that trumps ($\underline{a}$,T) or there is a sequence of extended agreements with strictly increasing utility for all members of T, whose utility converges to an upper bound. This sequence of extended agreements will have a subsequence

$$(\underline{d}(1),T),\ (\underline{d}(2),T),\ (\underline{d}(3),T)\ \ldots$$

such that

$\underline{d}^*(1), \underline{d}^*(2), \underline{d}^*(3), \ldots$

is a convergent sequence (recall the notational convention). Call the limit strategy vector $\underline{c}^*$.

Now choose a sequence of strategy vectors, one from each $\underline{d}(i)$, such that this sequence also converges to $\underline{c}^*$. Call this "diagonalized" sequence $\underline{c}$. Agreement $(\underline{c},T)$ trumps $(\underline{a},T)$ and no element of $\mathscr{C}$ trumps $(\underline{c},T)$. If $(\underline{c},T)$ is self-enforcing, we are done.

Suppose $(\underline{c},T)$ is not self enforcing. Then it is trumped by an optimal agreement $(\underline{f},V)$, with $V \subset T$. That means that all members of V strictly prefer $\underline{f}^*$ to $\underline{c}^*$, and that for some k

$$\underline{f}_i^k = \underline{c}_i^k \text{ for all } i \notin V$$

But for some $\underline{d}(m)$ and some k', $\underline{c}^k = \underline{d}^{k'}(m)$. And for all $i \notin V$, $\underline{f}_i^k = \underline{f}_i^{k'}$. That is to say, for k'

$$\underline{f}_i^{k'} = \underline{c}_i^{k'} \text{ for all } i \notin V.$$

And since all members of T prefer $\underline{c}^*$ to $\underline{d}^*(m)$, all members of V prefer $\underline{f}^*$ to $\underline{d}^*(m)$. That is to say, $(\underline{f},V)$ trumps $(\underline{d}^*(m),T)$, contradicting the assumption that all agreements in $\mathscr{C}$ were self-enforcing. ∎

Lemma A.4: An optimal extended agreement is not trumped by any self enforcing extended agreement.

Proof: Suppose otherwise that $(\underline{a},T)$ is optimal and trumped by a

self-enforcing ($\underline{b}$,S). Since one optimal extended agreement cannot be trumped by another, ($\underline{b}$,S) must not be optimal. But then by the preceding lemma there exists an optimal ($\underline{c}$,S) which trumps ($\underline{b}$,S). And by Lemma A.1, ($\underline{c}$,S) trumps ($\underline{a}$,T), a contradiction.■

<u>Lemma</u> <u>A</u> <u>5:</u> For any semi consistent partition, ($\underline{a}$,T) $\in$ G implies ($\underline{a}$,T) is optimal.

<u>Proof</u>: Identical to the proof for simple agreements in theorem 2 of the text.

<u>Lemma</u> <u>A</u> <u>6:</u> Suppose we have a semi-consistent partition with the property that for all coalitions T of size not exceeding k-1 ($\geq$ 2), ($\underline{a}$,T) $\in$ U implies that ($\underline{a}$,T) is self-enforcing. Then ($\underline{a}$,T) is optimal implies ($\underline{a}$,T) $\in$ G.

<u>Proof:</u> If ($\underline{a}$,T) is in U it is trumped by ($\underline{b}$,V) in U. Since the size of V is no greater than the size of T, ($\underline{b}$,V) is self enforcing by the hypothesis of the lemma. Then by lemma A.4, ($\underline{a}$,T) is not optimal.

If ($\underline{a}$,T) is in B, it is trumped by ($\underline{b}$,V) in G. By lemma A.5 ($\underline{b}$,V) is optimal, therefore ($\underline{a}$,T) is not optimal. ■

<u>Lemma</u> <u>A</u> <u>7:</u> Given a semi-consistent partition and a coalition T (with #T $\geq$ 2), ($\underline{a}$,T) $\in$ U implies that ($\underline{a}$,T) is self-enforcing.

<u>Proof</u>: Suppose #T = 2. Then ($\underline{a}$,T) not self-enforcing, means that there is a player i in T such that ($\underline{a}$,T) is trumped by an optimal ($\underline{c}$,{i}). For a single-player agreement to be optimal, it cannot be trumped by any agreement, and therefore ($\underline{c}$,{i}) $\in$ G, contradicting ($\underline{a}$,T) $\in$ U.

Now suppose the result is true for any T with #T $\leq$ k - 1, and consider

the case where $\#T = k$. If $(\underline{a},T) \in U$ is not self-enforcing then there exists an optimal $(\underline{b},W)$, with $W \subset T$, such that $(\underline{b},W)$ trumps $(\underline{a},T)$. Since $\#W \leq k - 1$, application of the preceding lemma implies that $(\underline{b},W)$ is in G, contradicting the assertion that $(\underline{a},T) \in U$. ∎

Lemma A 8: $(\underline{a},T)$ is optimal if and only if $(\underline{a},T) \in G$ in the semi-consistent partition.

Proof: Combining the two preceding lemmas we conclude that $(\underline{a},T)$ optimal implies $(\underline{a},T) \in G$. The reverse implication is lemma A.5. The uniqueness of the semi-consistent partition follows from the fact that the set of optimal agreements is constructed without any reference to consistent partitions. ∎

Lemma A 9: U is empty; i.e. the semi consistent partition is consistent.

Proof: Suppose $(\underline{a},S) \in U$. By lemma 3, $(\underline{a},S)$ is self-enforcing. By the preceding lemma, $(\underline{a},S) \notin G$ implies $(\underline{a},S)$ is not optimal. Thus there is $(\underline{c},S)$ which is optimal and which trumps $(\underline{a},S)$. By the preceding lemma $(\underline{c},S) \in G$ trumps $(\underline{a},S) \in U$, a contradiction. ∎

Proof of Corollary:

The equivalence follows from the following fact: $(\underline{a},T)$ is self enforcing if and only if it is not trumped by any self enforcing $(\underline{b},S)$ with $S \subset T$.

"If" follows from the fact that optimal agreements are a subset of self enforcing agreements. To prove "Only if" we show that if $(\underline{a},T)$ is trumped by a self enforcing $(\underline{b},S)$ with $S \subset T$, then it is trumped by an optimal $(\underline{c},S)$. This follows by applying lemma A.3 and lemma A.1.

References

Bernheim, D., B. Peleg, and M. Whinston, "Coalition-Proof Nash Equilibrium" Part I. Journal of Economic Theory, 1987.

Dutta, Bhaskar, Debraj Ray, Kunal Sengupta and Rajiv Vohra, "A Consistent Bargaining Set," Journal of Economic Theory, forthcoming.

Greenberg, Joseph "Stable Standards of Behavior: A unifying approach to solution concepts" Stanford University, IMSSS Economics Series, Technical Report 484, March, 1986.

Kahn, Charles and Dilip Mookherjee, "Decentralized Exchange and Efficiency Under Moral Hazard" working paper, University of Illinois, April 1989.

Roth, Alvin E., "Subsolutions and the Supercore of Cooperative Games," Mathematics of Operations Research, Vol. 1, No. 1, February, 1976, pp. 43-49.

Footnotes

[1]As shown below, the difficulties arise even with compact strategy spaces and continuous utility function.

[2]For the use of a similar generalization for cooperative games, see Dutta et al. [1989] in the case of the bargaining set, and Roth [1976] in the case of the core.

[3]It would be equivalent to give the definition with only one-way implications in the descriptions of G and B; for the consistent system the reverse implications are a consequence of the fact that {G,B} forms a partition of the set of all agreements. Note the obvious parallel between consistency and the von Neumann Morgenstern [1947] solution concept.

[4]In the definition of a semi-consistent system, unlike the definition of a consistent system, the double implications are crucial. It is possible that in a semi-consistent partition, the good set is empty, (and therefore the bad set is as well). This is the case in example 1 above. However the good set will not be empty if the payoff functions are continuous and individual players' strategy sets are closed.

[5]Again, having an open set of strategies is not necessary; more complicated examples can be derived in which payoffs are continuous on a compact strategy space.

[6] A natural extension of this approach will handle games with discontinuous payoffs by modifying (iii) in the definition of trumping as follows:

$$\liminf U_i(b^k) > \limsup U_i(a^k) .$$

To handle non-compact action spaces as well, it will be necessary to substitute monotonic sequences for convergent sequences of strategy vectors, and to use the overtaking criterion for comparison of limit payoffs.

[7] An alternate proof of uniqueness is provided in theorem 6.